

Developing Genome-Scale Models of Microbial Metabolism

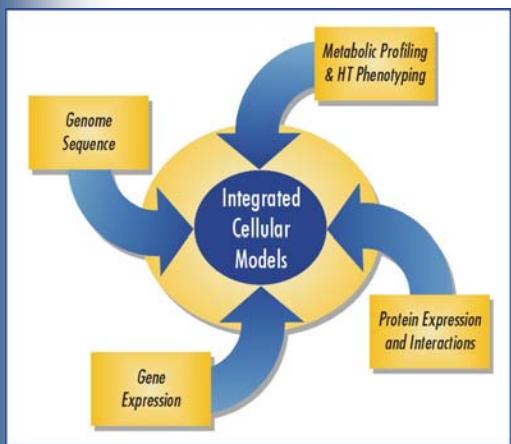
Christophe H. Schilling, Ph.D.
Chief Technical Officer
Genomatica, Inc.

Ninth DOE Genome Workshop
January 29, 2002

© 2002 Genomatica, Inc. All rights reserved.



Developing Microbial Models of Metabolism



General Topics

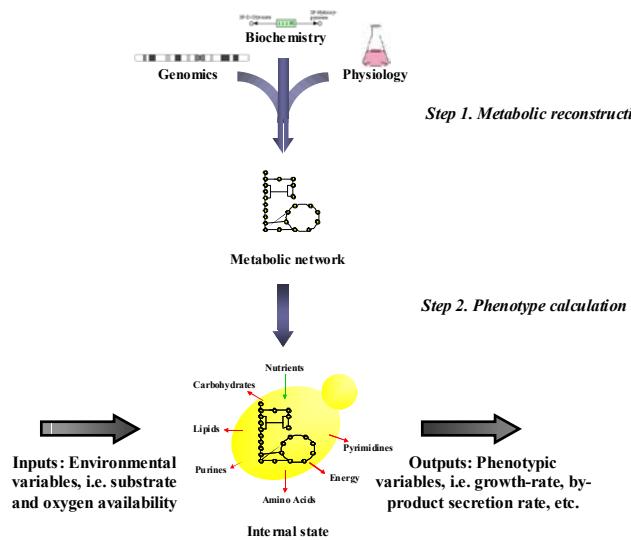
- Model development process
- Constraints-based modeling approach
- Genome scale microbial models
- Iterative model development
- Modeling technology development
- Model centric data integration (modeling & experimentation)

© 2002 Genomatica, Inc. All rights reserved.

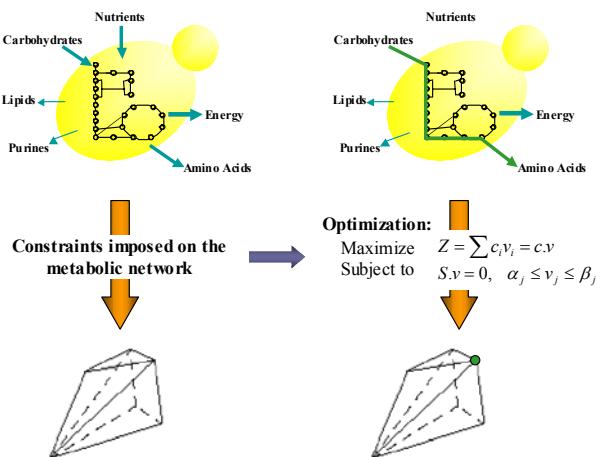
DOE Genome Workshop-1/29/02



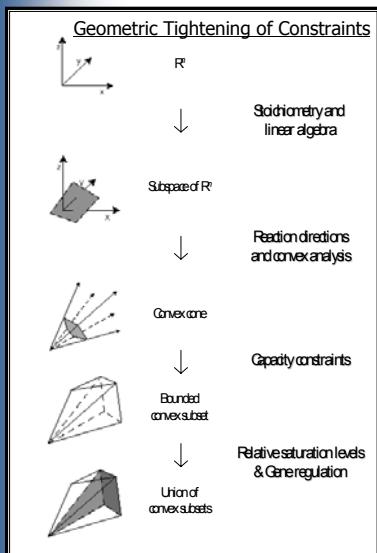
The Model Development Process



Calculating Phenotypes Using a Constraint-based Approach



Data Driven Constraints-Based Modeling



- Systematic imposition of constraints
- Determine what an organism is **not** capable of
- Determine optimal performance within constraints
- Types of constraints:
 - physico-chemical constraints (non-adjustable, hard)
 1. Stoichiometry of individual metabolic reactions
 2. Thermodynamics limiting the direction of reactions
 - system-specific constraints (adjustable)
 3. Maximum flux capacities
 4. kinetic constraints & time constants of reactions
 5. Genetic regulation & transcriptional control

© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Developed Microbial Models

Organism	<i>Escherichia</i>	<i>Haemophilus</i>	<i>Helicobacter</i>	<i>Saccharomyces</i>	<i>Bacillus</i>
	<i>coli</i>	<i>influenza</i>	<i>pylori</i>	<i>cerevisiae</i>	<i>subtilis</i>
Genome Characteristics					
Genome Size	4.6Mb	1.83Mb	1.66Mb	12.1Mb	4.2 Mb
Total Genes	4288	1703	1590	6281	4219
In Silico Model Characteristics					
Total Metabolic Genes	644	400	290	697	577
Percentage of Genome	15%	23%	18%	11%	14%
Metabolic Reactions	720	461	381	1212	783
Metabolites	436	343	340	734	599

DOE Microbial Cell Projects:

- *Geobacter sulfurreducens*
- *Shewanella oneidensis*

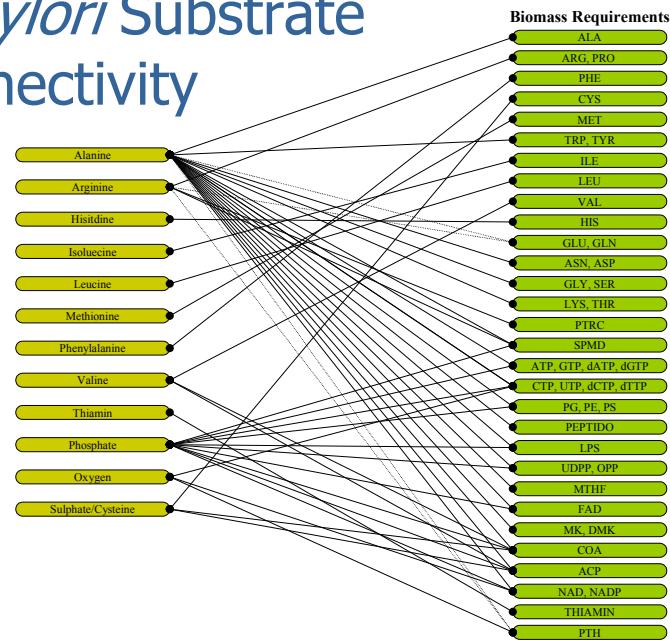
© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



H. pylori Substrate Connectivity

Helicobacter pylori
Minimal Substrate Requirements



© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



H. Pylori Deletion Analysis (comparison to in vitro essential gene study)

Deleted Reaction	Gene	Locus ID	In Vitro Essentiality*	In Silico Essentiality				
				Minimal	Minimal + Glucose	Minimal + Carbon Sources	Minimal + Carbon Sources + Amino Acids	Agreement
ACPS	acpS	HP0808	E	-	-	-	-	✓
AROQ	aroQ	HP1038	E	-	-	-	-	✓
DCD	dcd	HP0372	NE	+	+	+	+	✓
EDA	eda	HP1099	E	+	+	+	+	
EDD	edd	HP1100	E	+	+	+	+	
FBP	fbp	HP1385	E	-	+	+	+	
GLCD	glcD	HP0509	E	-	-	-	-	✓
GPT1/2/3	gpt	HP0735	E	+	+	+	+	
MQO	mqa	HP0086	E	-	-	+	+	
MURB	murB	HP1418	E	-	-	-	-	✓
MURF	murF	HP0740	E	-	-	-	-	✓
OOR	oorA, oorB, oorC, oorD	HP0589, HP0590, HP0591, HP0588	E	+	+	+	+	
POR	porA, porB, porC, porD	HP1110, HP1111, HP1108, HP1109	E	-	-	-	-	✓
PYRE	pyrE	HP1257	E	-	-	-	-	✓
SPEE	speE	HP0832	NE	+	+	+	+	✓
TAL	tal	HP1495	NE	+	+	+	+	✓
TRXB	trxA, trxB	HP0824, HP0825	E	-	-	-	-	✓

* Chalker et. al., 2001, *J Bacteriol* 183:1259-68.

© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Engineering *E. coli* for Ethanol Production

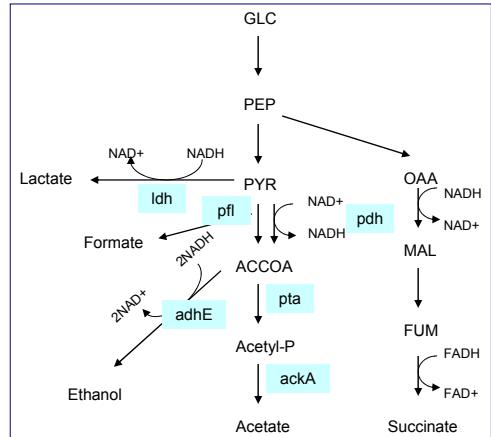
Model-dependent prediction:

mixed organic acid production under anaerobic conditions in the wild-type strain

Experimental Data Validation
Anaerobic Fermentation Condition

	In Silico	Yang, 1999	Chang, 1999
Glucose	1.00	1.00	1.00
Acetate	0.82	0.85	0.60
Ethanol	0.81	0.81	0.58
Lactate	0.00	0.05	0.26
Succinate	0.01	0.24	0.13
Pyruvate	0.00	0.00	0.03
Formate	1.73	1.66	1.13

All data in molar yields



© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Engineering *E. coli* for Ethanol Production

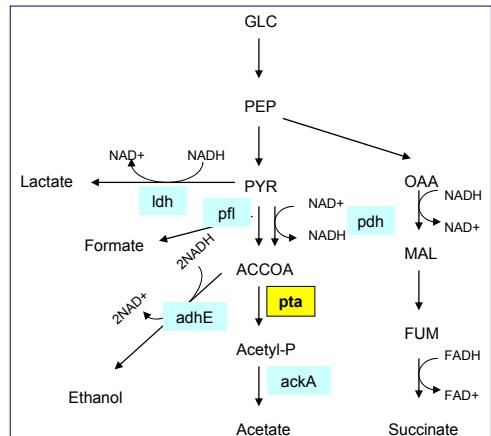
Model-dependent prediction:

***pta* knockout leads to increased lactate production (no ethanol increase)**

Experimental Data Validation
PTA Mutant Under Anaerobic Fermentation Condition

	In Silico	Yang, 1999	Chang, 1999
Glucose	1.00	1.00	1.00
Acetate	0.01	0.03	0.06
Ethanol	0.00	0.08	0.07
Lactate	1.45	1.54	1.52
Succinate	0.26	0.22	0.19
Pyruvate	0.00	0.05	0.02
Formate	0.00	0.10	0.08

All data in molar yields



© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Engineering *E. coli* for Ethanol Production

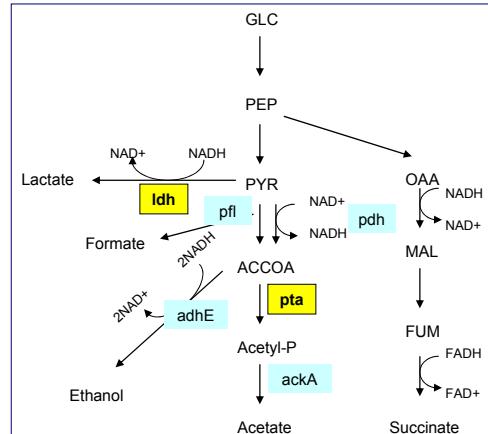
Model-dependent prediction:

***pta, ldh* double knockout leads to increased ethanol production**

Experimental Data Validation
PTA LDH Double Mutant Under Anaerobic Conditions

	In Silico	Yang, 1999	Chang, 1999
Glucose	1.00	1.00	1.00
Acetate	0.01	0.03	0.12
Ethanol	0.98	1.64	0.42
Lactate	0.00	0.01	0.00
Succinate	0.01	0.06	0.25
Pyruvate	0.68	0.12	0.36
Formate	1.07	1.67	0.44

All data in molar yields



© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02

Data Driven Constraints-Based Modeling

Engineering Design principles guide Cellular evolution and design

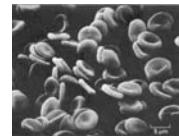
Engineering Design

- Objective
 - Build a bridge, Separation of proteins
- Constraints:
 - Geometry
 - Materials
 - Physical constants
 - Time & Money
- Design envelope
- Optimize design using free design variables



Biological Design

- Objective
 - Survival, Growth
- Constraints:
 - Max fluxes
 - Connectivity
 - P/C factors
 - Environment
- Solution space
- Optimize design using kinetic and regulatory variables

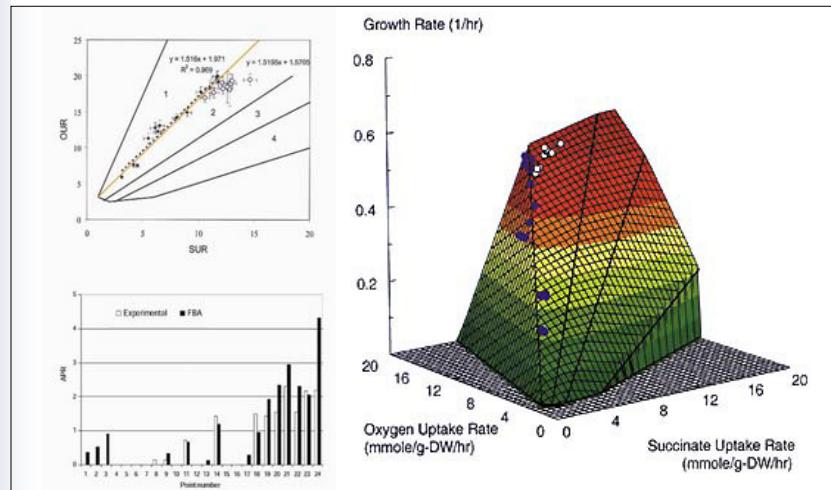


Cells are constrained in their behavior and seem to push close to these constraints ('life on the edge')

© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02

Optimal Growth of *E. coli* on Succinate Phenotypic Phase Plane Analysis



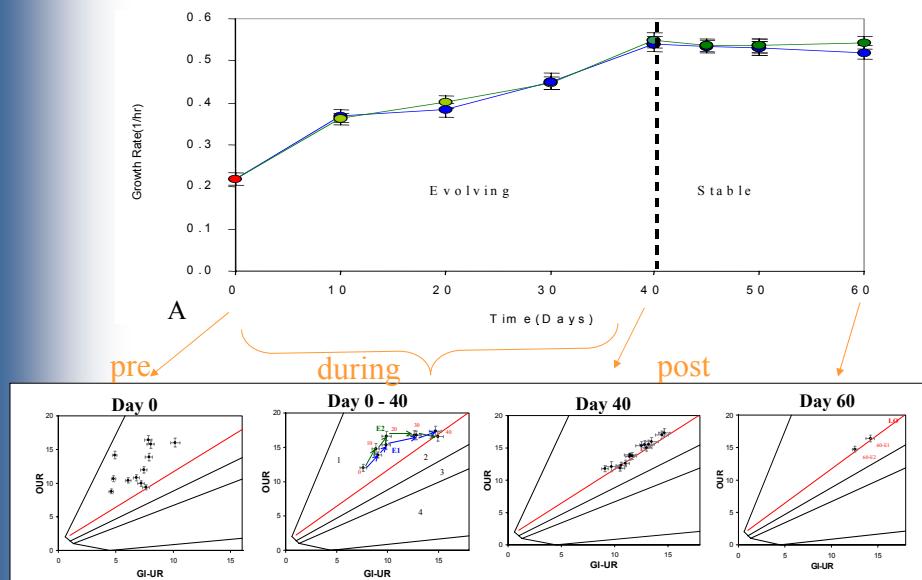
Edwards et. al., 2001, *Nature Biotechnol.* 19:125-130.

© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Cellular Evolution: Evolving Growth Rates on Glycerol

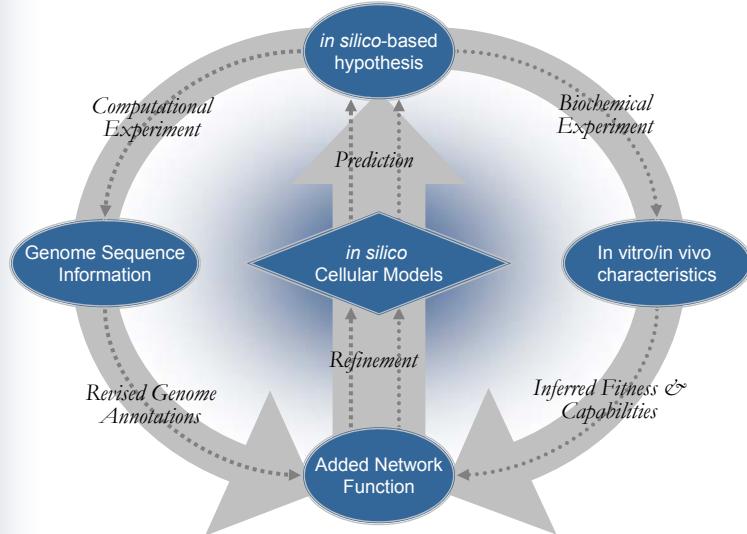


© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Iterative *In Silico* Model Building



© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Generations of Constraints-Based Models: Use of 'omics' Data

- 1st generation
 - Hard constraints
 - Determine capabilities (what)
- 2nd generation
 - Regulation of expression
 - Determine choices (why)
- 3rd generation
 - Regulation of activity
 - Determine trajectories (how)
- Genomics
 - Annotated sequence
 - Legacy data
- Expression profiling
 - Transcriptomics
 - Proteomics
- Concentration data
 - Metabolomics
 - Proteomics

© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Modeling Technology Development



- **Constraints-Based Methodologies** – continuous tightening of constraints based on new mathematical approaches (i.e. MILP, alternate objectives)



- **Signal Transduction/Gene Regulation Methods** – incorporation of regulatory interaction to control metabolic networks and development of supporting mathematics for 2nd generation models



- **Gene Expression Microarray Prediction** – *in silico* prediction of gene expression data sets and data integration for tighten model constraints

- **Metabolic Profiling Data Integration** – processing of metabolite concentration data to further tighten constraints

- **Pathway Analysis and Model Reduction** – extreme pathway calculations, flux cone volume calculations, and pathway reduction

- **Strain Evolution** – predicting and deciphering evolutionary principles guiding the development and optimization of metabolic performance

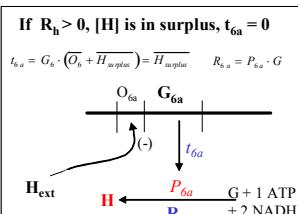
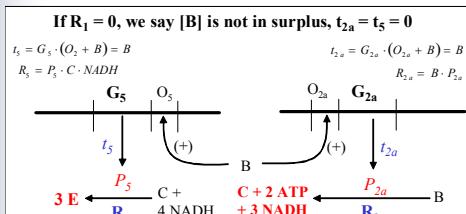
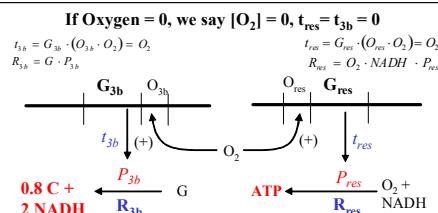
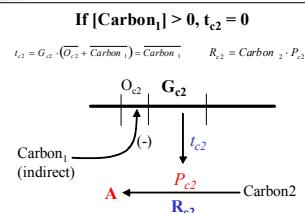
© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



Incorporation of Regulatory Structure: Condition-Dependent Flux Constraints

Hybrid Boolean/Stoichiometric Modeling

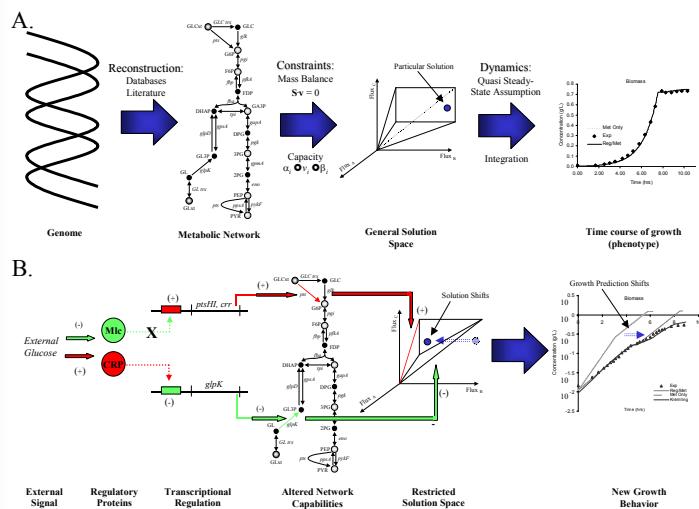


© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02

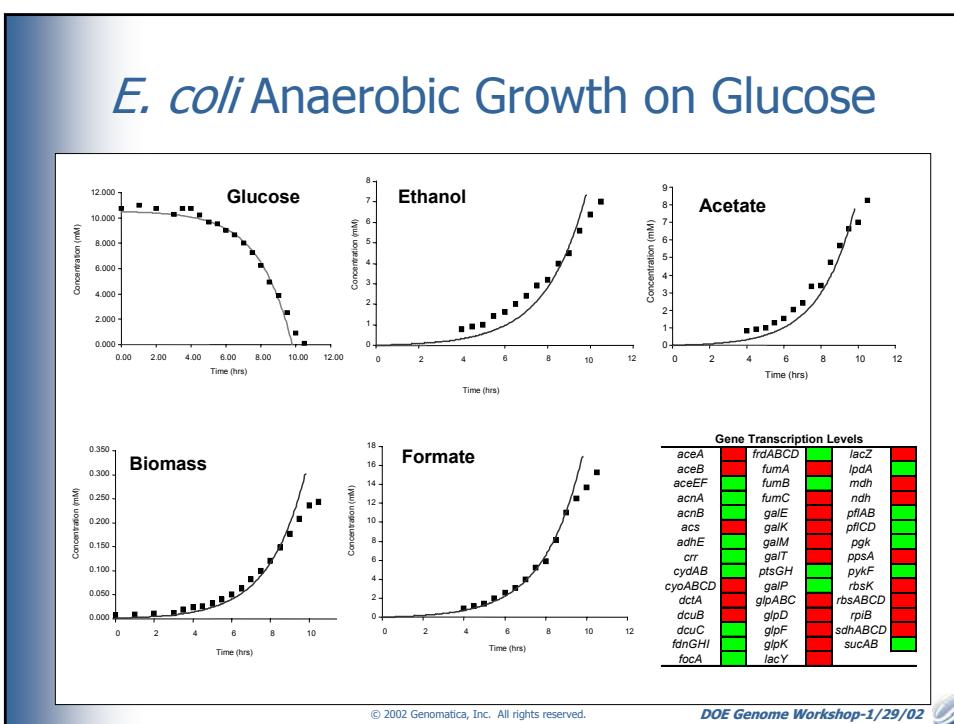


Combined Modeling of Metabolism and Transcriptional Regulation



© 2002 Genomatica, Inc. All rights reserved.

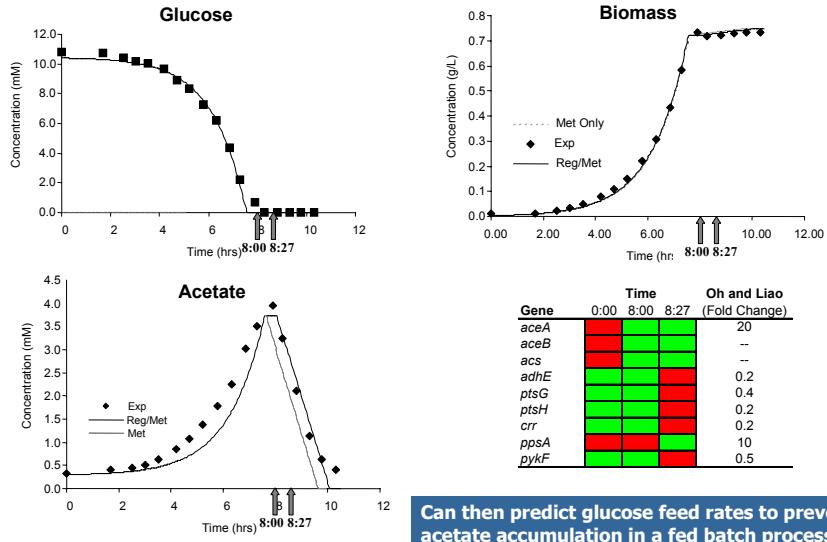
DOE Genome Workshop-1/29/02



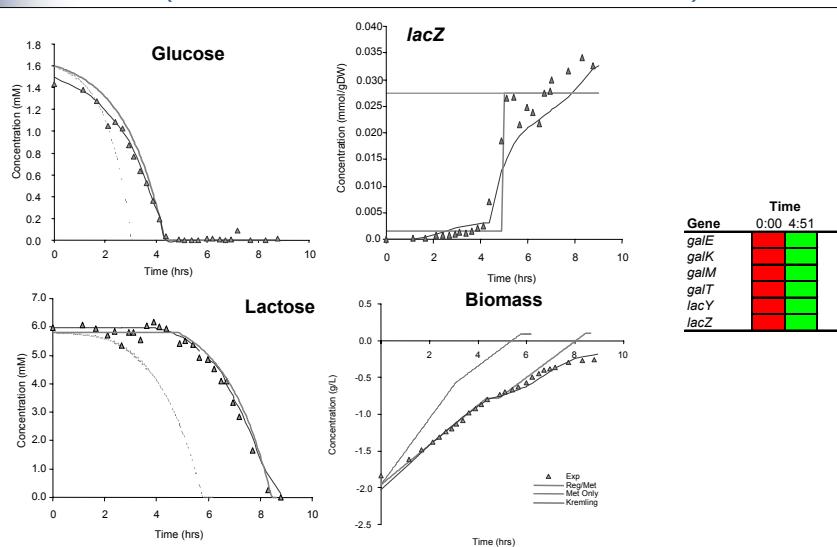
© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02

E. coli Aerobic Growth on Glucose with Acetate Reutilization



In Silico Prediction of Diauxic Response (Aerobic Growth on Glucose and Lactose in *E. coli*)



Summary

- Constraints-based modeling technology
 - Data-driven tightening of biological constraints
 - Genome-scale models of multiple microorganisms
 - Iterative model development process
- Developing Microbial Models for DOE Projects
 - *Geobacter sulfurreducens* – University of Massachusetts (Derek Lovley)
 - *Shewanella oneidensis* – Shewanella Federation
- Technology development supported by DOE
 - 2nd generation models to include regulation
 - MILP optimization approaches
 - Expanded experimental data integration

Acknowledgments

- Univ. of California, San Diego (Bernhard Palsson)
- Penn State University (Costas Maranas)
- Univ. of Delaware (Jeremy Edwards)

Funding Support

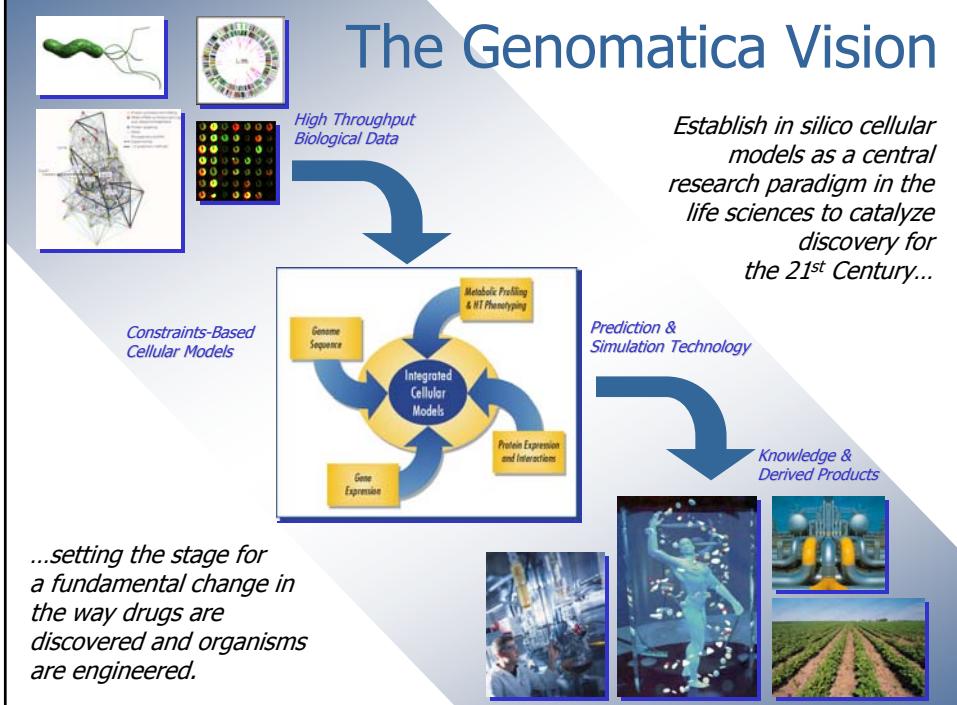
- Department of Energy (BER & ASCR)
- National Institutes of Health (NHGRI)

© 2002 Genomatica, Inc. All rights reserved.

DOE Genome Workshop-1/29/02



The Genomatica Vision



Developing Genome-Scale Models of Microbial Metabolism

Christophe H. Schilling, Ph.D.
Chief Technical Officer
Genomatica, Inc.

Ninth DOE Genome Workshop
January 29, 2002

© 2002 Genomatica, Inc. All rights reserved.

